HAL
open science

Défi d'annotation DEGELS2012 : la segmentation

Annelies Braffort, Leila Boutora

To cite this version:

Annelies Braffort, Leila Boutora. Défi d'annotation DEGELS2012 : la segmentation. JEP-TALN-RECITAL, ATALA/AFCP, Jun 2012, Grenoble, France. pp.1-8. hal-01803808

HAL Id: hal-01803808

https://hal.science/hal-01803808v1

Submitted on 1 Jun 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Défi d'annotation DEGELS2012 : la segmentation

Annelies Braffort *Leïla Boutora*
LIMSI-CNRS, Campus d'Orsay bat 508, BP133, 91403 Orsay cx
Laboratoire Parole et Langage, UMR 7039 CNRS/Aix-Marseille Univ, 13100 Aix-en-Provence
annelies.braffort@limsi.fr, leila.boutora@lpl-aix.fr

RÉSUMÉ

Dans cet article, nous présentons la deuxième édition du défi d'annotation de gestes et de langue des signes (DEGELS). Comme l'année dernière, l'objectif est d'organiser une campagne d'annotation dans le but de comparer des méthodologies d'annotation et d'analyse de corpus de gestes coverbaux en français oral et de langue des signes française (LSF) en soumettant aux chercheurs linguistes et informaticiens de ces domaines un corpus constitué pour l'occasion. L'édition 2012 se propose d'étudier les méthodes de segmentation des unités gestuelles, méthodes partagées par les communautés gestualiste et LSF. Après avoir présenté les objectifs, les enjeux scientifiques et le déroulement de cette manifestation scientifique, nous expliquons comment nous avons exploité les annotations réalisées par les cinq équipes participantes afin de préparer la journée de l'atelier.

ABSTRACT

DEGELS 2012 Annotation Challenge: Segmentation

In this paper, we present the second edition of the gesture and sign language annotation challenge (DEGELS). As last year, the goal is to organise an annotation campaign in order to compare methodologies for annotation and analysis of coverbal gestures in spoken French and French Sign Language (LSF) corpora. For that, we have submitted to linguists and computer scientists a corpus that has been specially created for this challenge. The DEGELS 2012 edition is dedicated to the study of segmentation of gestural units, a method shared by gestural and sign languages scientific communities. After presenting the objectives and the organisation of this event, we explain how we used the annotations provided by the five teams to prepare the workshop.

MOTS-CLÉS : Méthodologie d'annotation, schéma d'annotation, gestualité, multimodalité, Langue des Signes Française, segmentation

KEYWORDS : .Annotation methodology, annotation scheme, gesture, multimodality, French Sign Language, segmentation

1 Introduction

DEGELS (DEfi Geste Langue des Signes) est un atelier de comparaison d'annotation de corpus de gestes coverbaux et de langue des signes. L'objectif de cet atelier est de rassembler les communautés scientifiques étudiant la gestualité coverbale et la langue des signes autour des problématiques communes portant sur l'annotation de corpus.

Cet atelier prend la forme d'un défi d'annotation : nous fournissons aux équipes participantes des données vidéo extraites d'un corpus comparable de langue des signes et de français oral (voix et gestes). Les équipes doivent les annoter, fournir leur annotation ainsi qu'un article décrivant leur méthodologie (choix théoriques, schéma d'annotation, critères de choix...). Les organisateurs comparent les annotations entre elles sur une thématique différente chaque année et élabore un alignement des annotations lorsque c'est possible et une synthèse des différentes approches, les points communs, les différences. Cette synthèse est présentée lors de l'atelier et des échanges sont organisés autour de cas concrets issus des annotations des participants.

Le partage des corpus et des annotations passe par une bonne connaissance des méthodes et une description claire des schémas d'annotation utilisés ainsi que des critères appliqués (segmentation, choix parmi les catégories, etc.), au travers de guides d'annotation par exemple, tel que celui élaboré par Johnston (2011) pour l'annotation de la langue des signes australienne (AusLan). A travers cet atelier, notre objectif est d'initier la constitution de guides d'annotation pour la gestualité et la langue des signes.

Pour cette deuxième édition, nous proposons d'étudier les méthodes de segmentation en unités gestuelles et d'échanger sur les critères formels employés dans les études sur la LSF et en gestualité. Si un certain consensus existe dans la communauté des gestualistes (Kendon, 2004) (McNeill, 1992, 2005), ce n'est pas le cas pour les langues des signes où les approches peuvent être très différentes (Brentari & Wilbur, 2008) (Johnson & Liddell, 2011) (Hanke et al, 2011, 2012). Cependant, parmi elles, certaines peuvent se rapprocher de celle privilégiée en gestualité (Kita et al, 1998). Notre objectif est de mettre en regard ces différentes méthodes, avec leurs *pour* et leurs *contre*, afin d'expliciter les méthodes et les objectifs et les critères formels qui y sont associés puis de rédiger après l'atelier un document qui pourra servir de cadre pour les étudiants qui débutent, mais aussi les chercheurs qui n'ont pas toujours l'habitude de manipuler des critères formels.

2 Déroulement

L'atelier DEGELS 2012 a lieu le 8 juin 2012 à Grenoble, sur le campus de Saint-Martin d'Hères, dans le cadre de la conférence JEP-TALN 2012. L'interprétation en LSF y est assurée. Le corpus proposé aux participants, DEGELS1, est décrit dans (Boutora, Braffort, Bertrand 2011).

Les participants ont étudié les mêmes deux extraits de DEGELS1 qu'en 2011, de courte durée, dans l'optique d'aller plus loin dans l'explicitation des critères d'annotation, en particulier cette année les critères de segmentation. Pour les deux langues, les extraits portent sur la même portion d'itinéraire : la corniche en bord de mer qui mène du vieux

port aux plages de Marseille en passant par le David, statue emblématique d'un homme nu.

Le corpus, sous forme de fichiers vidéo et audio, était téléchargeable ainsi qu'un fichier texte comportant la traduction approchée en français écrit de l'extrait en LSF. Les fichiers vidéo sont proposés en différents formats de compression (xvid, IV5.1, Cinepak, non-compressé), afin d'assurer la compatibilité avec les logiciels d'annotation généralement employés (Elan, Anvil, et iLex). Pour le corpus de français oral, deux fichiers son sont fournis, correspondant à deux pistes (une par locuteur), séparées mais alignées entre elles et sur la vidéo.

L'ensemble du processus s'est déroulé en quatre étapes. Les participants ont eu accès aux corpus à partir du 13 février. Ils ont eu 8 semaines pour fournir leur annotation et leur article (4 avril). La notification a été envoyée aux auteurs le 23 avril. Les auteurs ont eu ensuite 2 semaines pour fournir la version finale de leur article (4 mai).

Cette année, 8 équipes se sont inscrites et 5 sont allées jusqu'au bout du processus. Ces équipes regroupent 11 intervenants, dont 7 gestualistes et 4 Langue des Signes. Les cinq équipes participantes sont constituées de la manière suivante :

Gestualité

- Équipe n°2 : G. Ferré (LLING Nantes)
- Équipe n°5 : D. Boutet (SFL Paris 8), K. Martel (PALM Caen) et M. Blondel (SFL Paris 8)
- Équipe n°8 : M. Tellier (LPL Aix-Marseille), B. Azaoui (DIPRALANG Montpellier 3) et J. Saubesty (LPL Aix-Marseille)

Langue des signes

- Équipe n°3 : I. Estève et A. Millet (LIDILEM Grenoble 3)
- Équipe n°7 : F. Lefebvre-Albaret (WebSourd) et J. Segouat (WebSourd)

3 Préparation de l'atelier

Le thème de cette année est la **segmentation** des unités gestuelles. Cette opération correspondant à un point de vue "bas-niveau" est souvent un préalable à l'étiquetage de l'élément annoté. Les annotations de DEGELS2011 ont montré une diversité des méthodes employées, pour lesquels une explicitation des critères de segmentation n'est pas toujours proposée. Il s'agit cette année d'expliciter ces critères, de les discuter, pour déboucher sur l'élaboration d'un guide d'annotation collectif qui reprendra ces critères, qui peuvent être différents en fonction de l'approche théorique ou méthodologique pratiquée.

Les éléments qu'il a été proposé d'étudier sont :

- Les éléments corporels , manuel ou non-manuel (uniquement les mains, d'autres articulateurs, une combinaison d'articulateurs...)
- Les temps de début et de fin des unités gestuelles

- Les critères formels utilisés pour définir les bornes de début et fin des unités gestuelles (manuel et non manuel, mouvement, forme de la main, direction du regard, autre...)
- Les phases temporelles des unités gestuelles (ex : *preparation, stroke, hold...*, ou *posture, transition, detention...* selon le modèle théorique)
- Les critères formels utilisés pour définir les bornes de début et fin de ces phases (mouvement, forme de la main, autre...)
- Les temps de début et de fin des phases gestuelles
- Les algorithmes employés en cas de segmentation automatique ou semi-automatique
- les méthodes de contrôle ou de validation de l'annotation le cas échéant.

Afin de pouvoir étudier en détail ces aspects le jour de l'atelier, il a été demandé aux participants d'inclure dans leur schéma d'annotation des pistes précises visant à segmenter les composantes manuelles et non-manuelles, chacune de ces pistes pouvant être affinée à l'aide de pistes filles :

Pour les composantes manuelles, chaque annotation devait comporter au minimum les trois pistes suivantes :

- "Segmentation MD" ou "Segmentation RH" (annotation de l'élément "main droite")
- "Segmentation MG" ou "Segmentation LH" (annotation de l'élément "main gauche")
- "Segmentation 2M" ou "Segmentation 2H" (annotation de l'élément "deux mains")

Pour les composantes non-manuelles, les pistes demandées étaient :

- "Segmentation Regard" ou "Segmentation Gaze" (annotation de la direction du regard)
- "Segmentation Buste" ou "Segmentation Chest" (annotation de l'élément "buste")
- "Segmentation Tête" ou "Segmentation Head" (annotation de l'élément "tête")

Chaque équipe a annoté soit l'extrait de corpus de gestualité, soit l'extrait de LSF. Toutes les équipes ont utilisé le logiciel Elan, sauf une qui a utilisé Anvil. Nous avons donc réalisé les alignements des annotations sous Elan, en exportant les annotations Anvil sous le format CSV qui peut être importé sous Elan.

Au vu des annotations, nous avons décidé d'aligner plus particulier les phases gestuelles, car ce niveau a été étudié par l'ensemble des équipes participantes. Ce niveau permet d'une part de visualiser la segmentation des unités gestuelles ainsi que des phases en elles-mêmes. Les figures 1 et 2 montrent un extrait des alignements effectués respectivement pour la gestualité et pour la LSF.

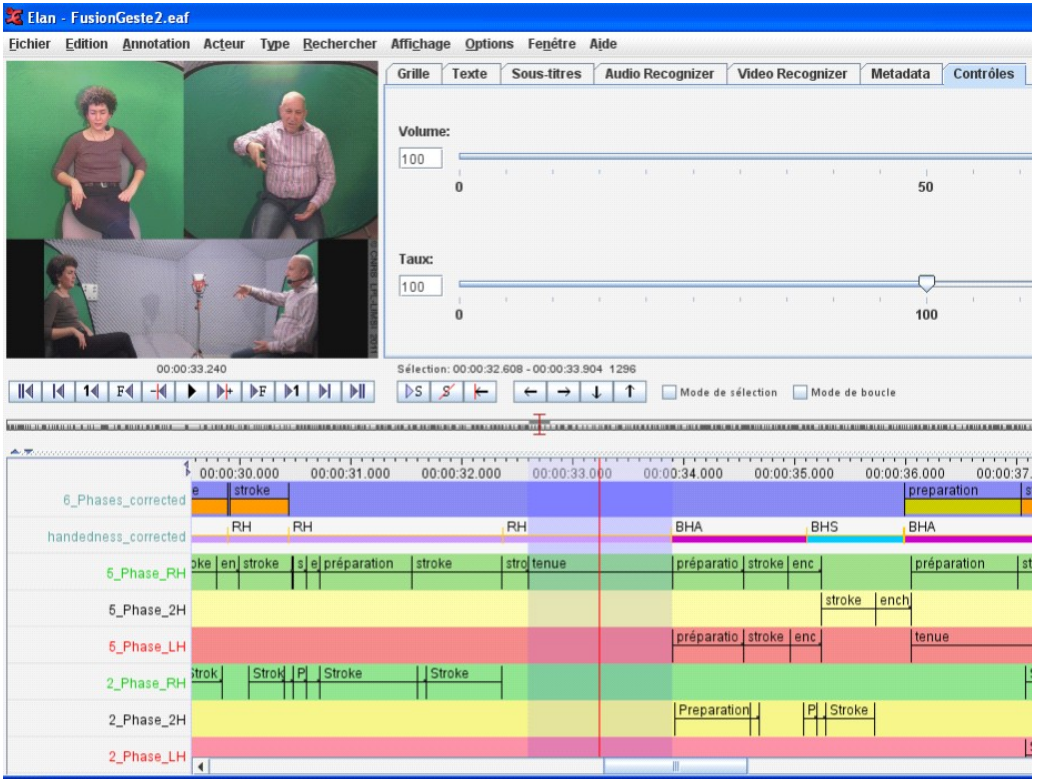

Figure 1 : Extrait de l'alignement des annotation pour la gestualité

Pour la gestualité (figure 1), l'équipe 6 a réalisé une étude sur l'accord inter-annotateur et a proposé une annotation résultant d'un accord après coup de trois annotateurs (1ère piste). Les autres équipes (5 et 2) ont utilisé trois pistes dédiées respectivement à l'annotation des événements observés pour la main droite (pistes vertes 3 et 6), la main gauche (pistes rouges 5 et 8) et les deux mains ensemble (pistes jaunes 4 et 7), tandis que l'équipe 6 a utilisé une seule piste pour décrire les événements manuels et une seconde piste (piste 2) pour indiquer s'il s'agit d'un geste monomanuel droite ou gauche ou d'un geste bimanuel symétrique ou non.

Pour la LSF (figure 2), les deux équipes ont utilisé trois pistes pour décrire les événements observés pour la main droite (pistes vertes 1 et 4), la main gauche (pistes rouges 3 et 6) et les gestes bimanuels (pistes jaunes 2 et 5). L'équipe 7 a en plus créé une piste avec des valeurs numériques représentées sous forme de courbes pour visualiser des mesures de vitesse des mains mesurées en annotant manuellement la position des mains dans les images constituant la vidéo.

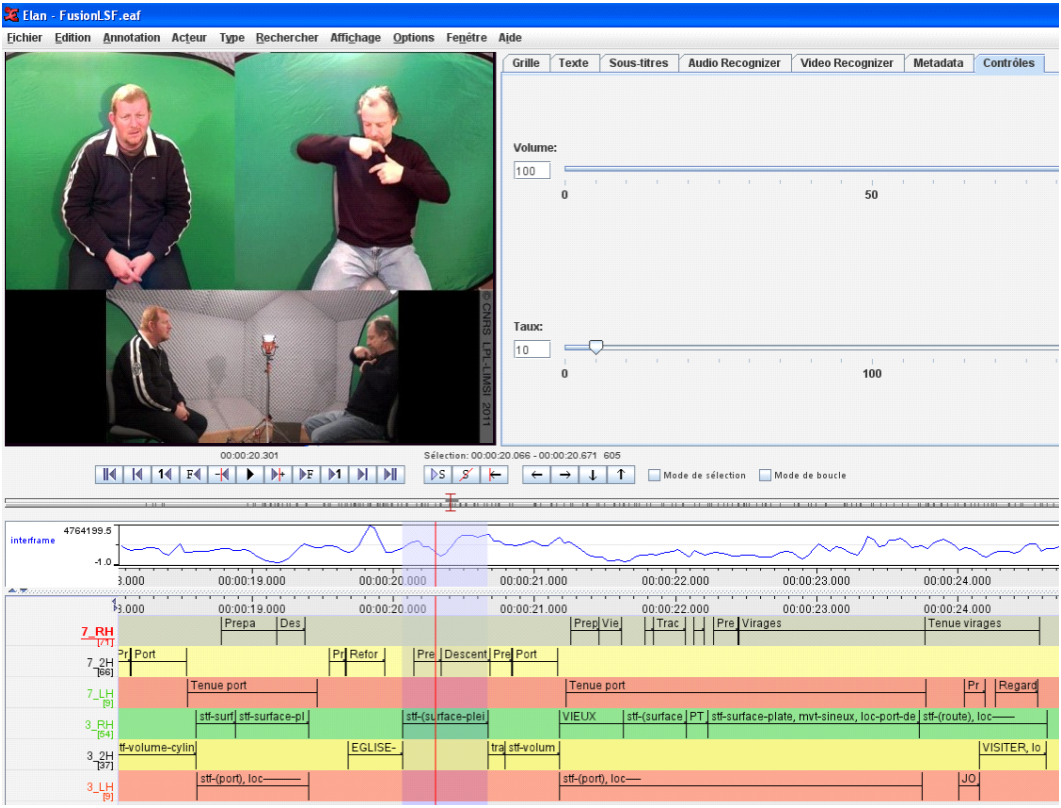

Figure 2 : Extrait de l'alignement des annotation pour la LSF

A partir de ces annotations, nous avons sélectionné quelques extraits qui nous ont semblé représentatifs pour les différences que l'on observe soit sur les choix de segmentation, soit sur les choix des pistes utilisées. Ces exemples seront débattus lors de la journée de l'atelier le 8 juin à Grenoble.

Nous avons par ailleurs invité deux présentations dédiées au traitement automatique qui peut être appliqué aux données pour l'aide à l'annotation, pour l'alignement texte/son d'une part (Bigi 2012), et pour la segmentation automatique de gestes par traitement d'images d'autre part (Gonzales Preciado 2012).

4 Conclusion

Dans cet article, nous avons présenté les objectifs et le déroulement de DEGELS 2012, deuxième édition d'un atelier qui consiste à comparer des méthodes d'annotation et d'analyse de corpus de gestes coverbaux en français oral et de LSF. L'édition de cette année s'est centrée sur la segmentation des unités gestuelles.

Pour chaque corpus (gestualité et LSF) nous avons aligné les annotations des phases des

événements manuels afin de pouvoir sélectionner quelques extraits à partir desquels lancer les discussions lors de la journée de l'atelier. Nous espérons ainsi pouvoir contribuer à l'élaboration d'un guide d'annotation à destination des étudiants et des chercheurs qui regrouperait les propositions issues de cet atelier.

L'enjeu de cet atelier est également de montrer l'impact des outils et des méthodologies sur l'analyse elle-même, pour prendre conscience de son existence et favoriser à terme des pratiques qui ne se laissent pas contraindre par les outils existants plus ou moins adaptés, mais au contraire de formuler les besoins des chercheurs afin d'aller vers le développement d'outils qui soutiennent réellement l'exploration linguistique.

Remerciements

Les auteurs remercient les relecteurs et les organisateurs de TALN et de DEFT pour leur aide à tous les niveaux, ainsi que les organisatrices locales de DEGELS 2012.

Cette manifestation a reçu le soutien financier du LIMSI et du LPL, ainsi que d' organismes de recherche (en cours de décision au moment de la rédaction de cet article) pour le financement de l'interprétariat en LSF.

Références

Corpus Degels1 oai:sldr.fr:crdo000767

Bigi, B. (2012). SPPAS : un outil « user-friendly » pour l'alignement texte/son. *In Actes de Degels 2012*, Grenoble, France.

Boutet, D., Martel, K., Blondel, M. (2012). Par où couper pour aller à la plage ? *In Actes de Degels 2012*, Grenoble, France.

Boutora, L., Braffort, A., Bertrand, R. (2011). Présentation et premiers résultats du défi d'annotation DEGELS2011 sur un corpus bilingue de français oral et de langue des signes. *In Actes de Degels 2011*, Montpellier, France.

Brentari, D., Wilbur, R. (2008). A cross-linguistic study of word segmentation in three sign languages, sign languages: spinning and unraveling the past, present and future. *Theoretical Issues in Sign Language Research Conference*, Florianopolis, Brazil, December 2006. Quadros (ed.). Editora Arara Azul. Petrópolis/RJ. Brazil.

Ferré, G. (2012). Critères de segmentation de la gestualité co-verbale. *In Actes de Degels 2012*, Grenoble, France.

Gonzales-Preciado, M. (2012). Un système de segmentation automatique de gestes appliqué à la Langue des Signes. *In Actes de Degels 2012*, Grenoble, France.

Hanke, T., Matthes, S., Regen, A., Storz, J., Worseck, S., Eliott, R., Glauert, J., Kennaway, R. (2011). Using timing information to improve the performance of avatars. *In Second International Workshop on Sign Language Translation and Avatar Technology (SLTAT)*,. Dundee, Scotland.

Hanke, T., Matthes, S., Regen A., Worseck S. (2012). Where Does a Sign Start and End?

Segmentation of Continuous Signing. *In Proceedings of LREC 2012 (Language Resources and Evaluation Conference) workshop RPSL (Representation and Processing of Sign Languages) : Interaction between Corpus and Lexicon.* Istanbul, Turkey.

Johnson, R. E., Liddell, S. K. (2011). A Segmental Framework for Representing Signs Phonetically. *Sign Language Studies* 11(3), pp.408-463. Washington: Gallaudet University Press

Johnston, T. (2011). Auslan Corpus Annotation Guidelines, November 2011. http://www.auslan.org.au/video/upload/attachments/AuslanCorpusAnnotationGuidelines30November2011.pdf

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

Kita, S., van Gijn, I., van der Hulst, H. (1998). Movement Phases in Signs and Co-speech Gestures, and Their Transcription by Human Coders. Gesture and Sign Language in Human-Computer Interaction, Lecture Notes in *Computer Science*, Vol. 1371/1998, 23-35.

Lefebvre-Albaret, F., Segouat, J. (2012). Influence de la segmentation temporelle sur la caractérisation de signes. *In Actes de Degels 2012*, Grenoble, France.

McNeill, D. (1992). *Hand and Mind: What gestures reveal about thought.* Chicago: The University of Chicago Press.

McNeill, D. (2005). *Gesture & thought.* Chicago: The University of Chicago Press.

Millet, A., Estève, I. (2012). Segmenter et annoter le discours d'un locuteur de LSF : permanence formelle et variabilité fonctionnelle des unités. *In Actes de Degels 2012*, Grenoble, France.

Tellier, M., Azaoui, B., Saubesty, J. (2012). Segmentation et annotation du geste : Méthodologie pour travailler en équipe. *In Actes de Degels 2012*, Grenoble, France.